## DETAILED ACTION

### *Claim Rejections - 35 USC § 102*

The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.
>
> (e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

Claims 19 and 39 are rejected under 35 U.S.C. 102(b) as being anticipated by ~~Nakatoh~~ et al. (US Patent No. ~~5611019~~).
Laudell 4811399

1. Regarding claims 19 and 39, Sewall et al. disclose a speech recognition system and a method comprising:

Means and a method for receiving an input signal representative of an audio signal (104 of figure 1);

an apparatus and method for detecting the presence of speech within the input signal (col. 6, ln. 1-3); and

recognition processing means and a method for performing a recognition processing of the portion of the input signal corresponding to speech (col. 6, ln. 55-66).

Claims 1-3, 7-8, 12-14, 18, 20-23, 27-28, 32-34, 38, and 40-42 are rejected under 35 U.S.C. 102(e) as being anticipated by Sewall et al. (US Patent No. 6708146).

2. Regarding claims 1 and 21, Sewall et al. disclose an apparatus and method for detecting the presence of speech within an input audio signal, comprising:

a memory for storing a predetermined function which gives, for a given set of audio signal values, a probability density for parameters of a predetermined speech model which is assumed to have generated the set of audio signal values, the probability density defining, for a given set of model parameter values, the probability that the predetermined speech model has those parameter values, given that the speech model is assumed to have generated the set of audio signal values (col. 13, ln. 13-36 and referring to figures 4-5);

means and a method for receiving a set of audio signal values representative of an input audio signal (element 10 of figures 4-5);

means and a method for applying the set of received audio signal values to stored function to give the probability density for model parameters for the set of received audio signal values (col. 15, ln. 6-25);

means and a method for processing the function with a set of received audio signal values applied to obtain values of parameters that are representative of the input audio signal (col. 15, ln. 6-25 or referring to figure 24); and

means and a method for detecting the presence of speech using obtained parameter values (col. 15, ln. 53-67).

3. Regarding claims 2 and 22, Sewall et al. further disclose that the processing means comprises means for drawing samples from probability density function and means for determining values of parameters that are representative of the speech from drawn samples (col. 16, ln. 16-30 and referring to figure 38).

4. Regarding claims 3 and 23, Sewall et al. further disclose that the drawing means is operable to draw samples iteratively from the probability density function (col. 16, ln. 11-43).

5. Regarding claims 7 and 27, Sewall et al. further disclose that the receiving means is operable to receive a sequence of sets of signal values representative of an input audio signal (col. 16, ln. 1-20) and wherein the applying means, processing means and detecting means are operable to perform their function with respect to each set of received audio signal values in order to determine whether or not each set of received signal values corresponds to speech (col. 13, ln. 13-36 and col. 15, ln. 6-25 or referring to figures 4-5).

6. Regarding claims 8 and 28, Sewall et al. further disclose that the processing means is operable to use the values of parameters obtained during the processing of a preceding set of signal values as initial estimates for the values of the corresponding

parameters of a current set of signal values being processed (col. 13, ln. 13-33 and col. 15, ln. 6-25).

103 obvious

7. Regarding claims 9 and 29, Sewall et al. further disclose a system that processes the sets of signal values in the sequence (col. 16, ln. 11-43 or figure 38), but fail to specifically disclose that the sets of signal values in the sequence are non-overlapping. However, it would have been obvious to one of ordinary skill in the art, at the time of invention, to incorporate non-overlapping frames to reduce processing time and power.

8. Regarding claims 12 and 32, Sewall et al. further disclose that the processing means is operable to vary the number of parameters used to represent the speech within the audio signal values (col. 11, ln. 11-17) and wherein the detecting means is operable to compare the number of parameters used to represent speech within the audio signal values with a predetermined threshold value, in order to detect the presence of speech within the audio signal (col. 8, ln. 56-67).

9. Regarding claims 13 and 33, Sewall et al. further disclose that the received speech signal values are representative of a speech signal generated by a speech source as distorted by a transmission channel between the speech source and the receiving means (element 10 of figures 4-5);

wherein the predetermined function includes a first part having first parameters which models the source (col. 13, ln. 13-36) and a second part having second parameters which models the channel (col. 11, ln. 11-17);

wherein the processing means is operable to obtain parameter values of at least the first parameters (col. 13, ln. 13-36); and

wherein the detecting means is operable to detect the presence of speech within the input audio signal from the obtained values of the first parameters (col. 15, ln. 53-67).

10. Regarding claims 14 and 34, Sewall et al. further disclose that the function is in terms of a set of raw speech signal values representative of speech generated by the source before being distorted by the transmission channel (col. 15, ln. 6-25), wherein the apparatus further comprises second processing means for processing the received set of signal values with initial estimates of the first and second parameters, to generate an estimate of the raw speech signal values corresponding to the received set of audio signal values (col. 13, ln. 60 to col. 14, ln. 9) and wherein the applying means is operable to apply the estimated set of raw speech signal values to the function in addition to the set of received signal values (col. 15, ln. 6-25).

11. Regarding claims 18 and 38, Sewall et al. further disclose that means for evaluating the probability density function for the set of received audio signal values using one or more derived samples of parameter values for different numbers of

parameter values, to determine respective probabilities that the predetermined speech model has those parameter values and wherein the processing means is operable to process at least some of the derived samples of parameter values and the evaluated probabilities to determine values of parameters that are representative of the audio speech signal (col. 13, ln. 13-37 and col. 15, ln. 6-25).

12. Regarding claims 20 and 40, Sewall et al. disclose a speech processing system and a method comprising:

means and a method for receiving an input audio signal (elements 10 of figures 4-5);

an apparatus a method for detecting the presence of speech within the input audio signal (col. 15, ln. 53-67); and

means and a method for processing the portion of the input audio signal corresponding to speech (col. 16, ln. 1-43 or figure 38).

13. Regarding claim 41, Sewall et al. disclose a computer readable medium storing computer executable process steps to cause a programmable computer apparatus to perform the method of claim 21 (col. 16, ln. 1-10).

14. Regarding claim 42, Sewall et al. disclose a processor implementable process steps for causing a programmable computing device to perform the method according to claim 21 (col. 16, ln. 53-55).

### *Claim Rejections - 35 USC § 103*

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

Claims 10-11, 15-17, 30-31, and 35-37 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sewall et al. (US Patent No. 6708146) in view of Burnett et al. (US Patent No. 6377919).

15. Regarding claims 10 and 30, Sewall et al. further disclose that the that the detecting means is operable to compare the value of at least one of the speech model coefficients with a prestored threshold value speech model (col. 13, ln. 1-10), but fail to specifically disclose that the speech model comprises an auto-regressive process model, wherein parameters include auto-regressive model coefficients. However, Burnett et al. teach that the speech model comprises an auto-regressive process model, wherein parameters include auto-regressive model coefficients (col. 11, ln. 1-3). The advantage of using the teaching of Burnett et al. in Sewall et al. is to minimize the level of impulsive noise.

Since Sewall et al. and Burnett et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the

art, at the time of invention, to modify Sewall et al. by incorporating the teaching of Burnett et al. in order to minimize the level of impulsive noise.

16. Regarding claims 11 and 31, Sewall et al. further disclose that the detecting means is operable to compare the values of a plurality of speech model coefficients with a corresponding plurality of predetermined values (). With the modification of Sewall et al. as discussed in claims 10 and 30, respectively, the speech model coefficients are the auto-regressive model coefficients. Thus, all method steps applied to the speech model coefficients would obviously apply to the auto-regressive model coefficients.

17. Regarding claims 15 and 35, Sewall et al. fail to disclose that the second processing means comprises a simulation smoother. However, Burnett et al. teach that the second processing means comprises a simulation smoother (col. 4, ln. 4-12). The advantage of using the teaching of Burnett et al. in Sewall et al. is to remove impulsive noise.

Since Sewall et al. and Burnett et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to modify Sewall et al. incorporating the teaching of Burnett et al. in order to remove impulsive noise.

18. Regarding claims 16 and 36, Sewall et al. fail to disclose that the second processing means comprises a Kalman filter. However, Burnett et al. teach that the

second processing means comprises a Kalman filter (col. 4, ln. 1-15). The advantage of using the teaching of Burnett et al. in Sewall et al. is to remove signal that don't have expected behaviors in time or frequency domains to make the system more efficient.

Since Sewall et al. and Burnett et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to modify Sewall et al. by incorporating the teaching of Burnett et al. in order to remove signal that don't have expected behaviors in time or frequency domains to make the system more efficient.

19. Regarding claims 17 and 37, Sewall et al. fail to disclose that the second part is a moving average model and wherein said second parameters comprise moving average model coefficients. However, Burnett et al. teach that the second part is a moving average model and wherein said second parameters comprise moving average model coefficients (col. 11, ln. 1-4). The advantage of using the teaching of Burnett et al. in Sewall et al. is to enable calculation of accurate transfer-function-filters.

Since Sewall et al. and Burnett et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to modify Sewall et al. by incorporating the teaching of Burnett et al. in order to enable calculation of accurate transfer-function-filters.

Claims 4 and 24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sewall et al. (US Patent No. 6708146) in view of Lennon et al. (US Patent No. 5116090).

20. Regarding claims 4 and 24, Sewall et al. fail to disclose that the processing means comprises a Gibbs sampler. However, Lennon et al. teach that the processing means comprises a Gibbs sampler (col. 12, ln. 18-41). The advantage of using the teaching of Lennon et al. in Sewall et al. is to obtain the optimum region label configuration for the frame.

Since Sewall et al. and Lennon et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to modify Sewall et al. by incorporating the teaching of Lennon et al. in order to obtain the optimum region label configuration for the frame.

Claims 5-6 and 25-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sewall et al. (US Patent No. 6708146) in view of Komissarchik et al. (US Patent No. 5799276).

21. Regarding claims 5 and 25, Sewall et al. fail to disclose that the processing means is operable to determine a histogram of drawn samples and wherein values of parameters are determined from the histogram. However, Komissarchik et al. teach that the processing means is operable to determine a histogram of drawn samples and

wherein values of parameters are determined from the histogram (col. 14, ln. 6-17). The advantage of using the teaching of Komissarchik et al. in Sewall et al. is to determine the likelihood weights of input speech model to increase accuracy.

Since Sewall et al. and Komissarchik et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to further Sewall et al. by incorporating the teaching of Komissarchik et al. in order to determine the likelihood weights of input speech model to increase accuracy.

+ Eatwell

22. Regarding claims 6 and 26, Sewall et al. fail to disclose that the processing means is operable to determine values of parameters using a weighted sum of drawn samples, and wherein the weighting is determined from the histogram. However, Komissarchik et al. further teach that the weighting is determined form the histogram (col. 5, ln. 6-17). The advantage of using the teaching of Komissarchik et al. in Sewall et al. is to enhance the detection accuracy.

Since Sewall et al. and Komissarchik et al. are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to modify Sewall et al. by incorporating the teaching of Komissarchik et al. in order to enhance the recognition accuracy.

The modified Sewall et al. still fail to specifically disclose that the processing means is operable to determine values of parameters using a weighted sum of drawn samples. However, Eatwell teaches that the processing means is operable to

determine values of parameters using a weighted sum of drawn samples (col. 5, ln. 1-10). The advantage of using the teaching of Eatwell in the modified Sewall et al. is to reduce noise in the signal.

Since the modified Sewall et al. and Eatwell are analogous art because they are from the same field of endeavors, it would have been obvious to one of ordinary skill in the art, at the time of invention, to further modify Sewall et al. by incorporating the teaching of Eatwell in order to reduce noise in the signal.

## *Conclusion*

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Huyen Vo whose telephone number is 703-305-8665. The examiner can normally be reached on M-F, 9-5:30.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Doris To can be reached on 703-305-4827. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Examiner Huyen X. Vo March 24, 2004

***